

Systematic Error-Correcting Codes for Rank Modulation

Hongchao Zhou, *Member, IEEE*, Moshe Schwartz, *Senior Member, IEEE*,
Anxiao (Andrew) Jiang, *Senior Member, IEEE*, and Jehoshua Bruck *Fellow, IEEE*

Abstract—The rank-modulation scheme has been recently proposed for efficiently storing data in nonvolatile memories. Error-correcting codes are essential for rank modulation, however, existing results have been limited. In this work we explore a new approach, *systematic error-correcting codes for rank modulation*. Systematic codes have the benefits of enabling efficient information retrieval and potentially supporting more efficient encoding and decoding procedures. We study systematic codes for rank modulation under Kendall's τ -metric as well as under the ℓ_∞ -metric.

In Kendall's τ -metric we present $[k+2, k, 3]$ -systematic codes for correcting one error, which have optimal rates, unless systematic perfect codes exist. We also study the design of multi-error-correcting codes, and provide two explicit constructions, one resulting in $[n+1, k+1, 2t+2]$ systematic codes with redundancy at most $2t+1$. We use non-constructive arguments to show the existence of $[n, k, n-k]$ -systematic codes for general parameters. Furthermore, we prove that for rank modulation, systematic codes achieve the same capacity as general error-correcting codes.

Finally, in the ℓ_∞ -metric we construct two $[n, k, d]$ systematic multi-error-correcting codes, the first for the case of $d = O(1)$, and the second for $d = \Theta(n)$. In the latter case, the codes have the same asymptotic rate as the best codes currently known in this metric.

Index Terms—flash memory, rank modulation, error-correcting codes, permutations, metric embeddings, Kendall's τ -metric, ℓ_∞ -metric, systematic codes

I. INTRODUCTION

THE rank-modulation scheme has been recently proposed for storing data efficiently and robustly in nonvolatile memories (NVMs) [8]. Its applications include flash memories [4], which are currently the most widely used family of NVMs, and several emerging NVM technologies, such as phase-change memories [2]. The rank-modulation scheme uses the relative order of cell levels to represent data, where a

cell level denotes a floating-gate cell's threshold voltage for flash memories and denotes a cell's electrical resistance for resistive memories (such as phase-change memories). Consider n memory cells, where for $i \in [n] = \{1, 2, \dots, n\}$, let $c_i \in \mathbb{R}$ denote the level of the i th cell. It is assumed that no two cells have the exact same level, which is easy to realize in practice. Let S_n denote the set of all $n!$ permutations over $[n]$. The n cell levels induce a permutation $[f_1, f_2, \dots, f_n] \in S_n$, where $c_{f_1} > c_{f_2} > \dots > c_{f_n}$. The rank-modulation scheme uses such permutations to represent data. It enables memory cells to be programmed efficiently and robustly, from lower levels to higher levels, without the risk of over-programming. It also makes it easier to adjust cell levels when noise appears without erasing cells, and makes the stored data more robust to asymmetric errors that change cell levels in the same direction [8], [9], [23].

Error-correcting codes are essential for data reliability. Intuitively, an error-correcting code is a set of elements in a metric space, no two of which are too close together under its distance measure. In the case of rank modulation, the space is S_n . As for the distance measure, it is usually chosen in such a way that small (common) errors in the physical medium correspond to a small distance in the metric space. In the context of rank modulation for NVMs, the two most studied distance functions are Kendall's τ -distance, and the ℓ_∞ -distance. It was suggested in [9] that small charge-constrained errors correspond to a small distance in Kendall's τ -metric. In contrast, in [23] it was shown that small limited-magnitude errors correspond to a small ℓ_∞ -distance.

Some results are known on error-correcting codes for rank modulation equipped with Kendall's τ -distance. In [9], a one-error-correcting code is constructed based on metric embedding, whose size is provably within half of the optimal size. In [1], the capacity of rank modulation codes is derived for the full range of minimum distance between codewords, and the existence of codes whose sizes are within a constant factor of the sphere-packing bound for any fixed number of errors is shown. Some explicit constructions of error-correcting codes have been proposed and analyzed in [18]. We also mention that the Ulam metric has been suggested as a generalization of Kendall's τ -metric and was recently studied in the context of error-correcting codes in [6].

There has also been some work on error-correcting codes for rank modulation equipped with the ℓ_∞ -distance. In [13], [23] some general constructions and bounds were given. A relabeling scheme, improving the distance of codes was suggested in [24]. Several counting problems, mainly concerning

This work was supported in part by a United States - Israel Binational Science Foundation (BSF) grant 2010075, a National Science Foundation (NSF) CAREER Award CCF-0747415, an NSF grant CCF-1217944, and an NSF grant CCF-1218005.

The material in this paper was presented in part at the 2012 IEEE International Symposium on Information Theory.

Hongchao Zhou is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA (e-mail: hongchao@mit.edu).

Moshe Schwartz is with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel (e-mail: schwartz@ee.bgu.ac.il).

Anxiao (Andrew) Jiang is with the Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843-3112, U.S.A. (e-mail: ajiang@cse.tamu.edu).

Jehoshua Bruck is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA (e-mail: bruck@paradise.caltech.edu).

ball size under the ℓ_∞ -metric, and optimal anticodes, were studied in [11], [12], [21], [22].

In this paper, we study *systematic* error-correcting codes for rank modulation as a new approach for code design. In the more common error-correcting setting over vectors equipped with the Hamming distance function, an $[n, k]$ -systematic code is a subset of length- n vectors whose projection onto a given set of k coordinates has all possible length- k vectors appearing exactly once. These k positions are referred to as the *information symbols*, whereas the rest of the positions are called *redundancy symbols*. If the code is linear, it is well known (for example, see [16]) that any code has an equivalent code with the same parameters that is also systematic. We shall be interested in the analog of systematic codes in the space of permutation with either Kendall's τ -distance or the ℓ_∞ -distance. Compared with the existing constructions of error-correcting codes for rank modulation, systematic codes have the benefit that they support efficient data retrieval, because when there is no error (or when error correction is not required), data may be retrieved by only reading the information symbols. Since every permutation induced by the information symbols appears in exactly one codeword, the codewords can be mapped efficiently to data (and vice versa) via enumerative source coding (e.g., by ordering permutations lexicographically) [5], [17]. In addition, the encoding algorithm of the error-correcting code can potentially be made very efficient by defining the positions of the redundant cells in the permutation as a function of the corresponding positions of the information cells.

We study the design of systematic codes, and analyze their performance. In Kendall's τ -metric we present families of $[k + 2, k]$ -systematic codes for correcting a single error. We show that they have optimal parameters among systematic codes, unless *perfect* systematic one-error-correcting codes, which meet the sphere-packing bound, exist. We also study the design of systematic codes that correct multiple errors, prove the existence of codes with minimum distance $n - k$ for any $2 \leq k < n$, as well as give a construction for a wide range of parameters. Furthermore, we prove that systematic codes have the same capacity as general error-correcting codes. This result establishes that, asymptotically, systematic codes are as strong in their error-correction capability as general codes. We also consider the ℓ_∞ -metric, and provide a general construction for systematic code whose asymptotic rate equals that of the best codes currently known.

The rest of the paper is organized as follows. In Section II we provide the basic notation and definitions used throughout the paper. In Section III we study systematic codes in Kendall's τ -metric. We turn in Section IV to explore systematic codes in the ℓ_∞ -metric. We conclude in Section V and present some open problems.

II. NOTATION AND DEFINITIONS

Let $[n] = \{1, 2, \dots, n\}$, and S_n denote the set of permutations over $[n]$. A permutation $f \in S_n$ is represented in single-line notation by $f = [f_1, f_2, \dots, f_n]$, where $f(i) = f_i$. We also denote the identity permutation by $\text{Id} = [1, 2, \dots, n]$. Finally,

we denote by f^{-1} the inverse of the permutation f , i.e., the permutation sending $f(i)$ to i .

Consider a metric over the permutations S_n with a distance function $d : S_n \times S_n \rightarrow \mathbb{N} \cup \{0\}$. An (n, M, d_{\min}) -code is a subset $C \subseteq S_n$ such that $|C| = M$, and $d(f, g) \geq d_{\min}$ for all $f, g \in C$, $f \neq g$. We say M is the *size* of the code, and d_{\min} is the minimum distance of the code.

In this work we shall consider two distance functions: Kendall's τ -distance, and the ℓ_∞ -distance. The latter ℓ_∞ -distance function is easily defined for all $f, g \in S_n$ by

$$d_\infty(f, g) = \max \{|f(i) - g(i)| \mid i \in [n]\}.$$

For the former, Kendall's τ -distance function, assume $f \in S_n$ is some permutation. An *adjacent transposition* on f switches the values of $f(i)$ and $f(i + 1)$ for some $i \in [n - 1]$. Kendall's τ -distance [10] between f and g , denoted $d_K(f, g)$, is defined as the minimal number of adjacent transpositions required to transform f into g . This is sometimes also called the *bubble-sort distance*.

We recall that in the rank-modulation scheme we have n memory cells labeled by $[n]$, and the level of the i th cell is denoted by $c_i \in \mathbb{R}$. Assume $c_{i_1} > c_{i_2} > \dots > c_{i_n}$, then the permutation stored by the n cells is $[i_1, i_2, \dots, i_n] \in S_n$ (see [8]). Assume a permutation $f \in S_n$ was stored, but a distorted version of it, $g \in S_n$, was eventually read. It was noted in [9] that small charge-constrained errors translate to small Kendall's τ -distance, denoted $d_K(f, g)$. In contrast, it was suggested in [23], [24], that limited-magnitude errors translate to small ℓ_∞ -distance on the inverse permutation, denoted $d_\infty(f^{-1}, g^{-1})$. This difference between storing the permutation or its inverse will play a role in defining two versions of systematic codes.

In order to define systematic codes we need to define two types of projections. Let $A = \{a_1, a_2, \dots, a_m\} \subseteq [n]$ be any subset, $a_1 < a_2 < \dots < a_m$. For any permutation $f \in S_n$, we define $f|_A$ to be the permutation in S_m that preserves the relative order of the sequence $f(a_1), f(a_2), \dots, f(a_m)$. Intuitively, to compute $f|_A$ we keep only the *coordinates* of f that appear in A , and then relabel the entries to $[m]$ while the keeping relative order. In a similar fashion we define

$$f|_A = (f^{-1}|_A)^{-1}.$$

Intuitively, to calculate $f|_A$ we keep only the *values* of f from A , and then relabel the entries to $[m]$ while keeping relative order.

Example 1. Let $n = 6$ and consider the permutation

$$f = [6, 1, 3, 5, 2, 4] \in S_6.$$

We take $A = \{3, 5, 6\}$. We then have

$$f|_A = [2, 1, 3],$$

since we keep positions 3, 5, and 6, of f , giving us $[3, 2, 4]$, and then relabel these to get $[2, 1, 3]$.

Similarly, we have

$$f|_A = [3, 1, 2],$$

since we keep the values 3, 5, and 6, of f , giving us $[6, 3, 5]$, and then relabel these to get $[3, 1, 2]$. \square

We are now in a position to define systematic codes in two different ways, depending on the metric.

Definition 2. An $[n, k, d]$ systematic code, C , for Kendall's τ -metric, is an $(n, k!, d)$ code such that

$$\left\{ f|^{[k]} \mid f \in C \right\} = S_k.$$

We call $[k]$ the information symbols of the code, and $\{k+1, k+2, \dots, n\}$ the redundancy symbols of the code.

Intuitively, if we have an $[n, k, d]$ -systematic code in Kendall's τ -metric, reading just the levels of the first k cells and comparing them, enables us to ascertain the relative positions of the values $1, 2, \dots, k$ in the stored permutation, and there is a unique codeword with this relative ordering. More precisely, assume a codeword $f \in C$ is stored. If the levels we read from the first k cells are c_1, c_2, \dots, c_k , and their ordering is $c_{i_1} > c_{i_2} > \dots > c_{i_k}$, then i_1 appears before i_2 in the codeword, appearing before i_3 , and so on, until i_k , i.e., $f^{-1}(i_1) < f^{-1}(i_2) < \dots < f^{-1}(i_k)$.

In contrast, in the setting of limited-magnitude errors and the ℓ_∞ -metric [23], the *inverse* of the permutation read from the cells is protected by an error-correcting code. Thus, if g is the codeword we want to store, we would physically write its inverse g^{-1} to the cells using the rank-modulation scheme. Then, reading just the levels of the first k cells, c_1, c_2, \dots, c_k , gives us the relative ordering of $g(1), g(2), \dots, g(k)$. This motivates the following definition.

Definition 3. An $[n, k, d]$ systematic code, C , for the ℓ_∞ -metric, is an $(n, k!, d)$ code such that

$$\left\{ g|_{[k]} \mid g \in C \right\} = S_k.$$

We call $[k]$ the information coordinates of the code, and $\{k+1, k+2, \dots, n\}$ the redundancy coordinates of the code.

III. SYSTEMATIC CODES IN KENDALL'S τ -METRIC

This section is devoted to the study of systematic codes in Kendall's τ -metric. In Section III-A we introduce further notation and some useful lemmas. In Section III-B we study systematic one-error-correcting codes. We turn, in Section III-C, to the case of general systematic error-correcting codes. Finally, in Section III-D, we analyze the capacity of systematic codes.

A. Preliminaries

We let \mathbb{Z}_n denote the set of integers $\{0, 1, \dots, n-1\}$, as well as the additive group over these integers with addition modulo n . It is well known (see [9], and references therein) that there is a one-to-one correspondence between the permutations of S_n and *factoradic* representations, which are mixed-radix vectors from

$$\mathbb{Z}_n! = \mathbb{Z}_1 \times \mathbb{Z}_2 \times \dots \times \mathbb{Z}_{n-1} \times \mathbb{Z}_n.$$

Let $f \in S_n$ be any permutation. The factoradic representation corresponding to f is a vector $v = [v_1, \dots, v_n] \in \mathbb{Z}_n!$ such that $v_i \in \mathbb{Z}_i$ equals

$$v_i = \left| \left\{ j \mid j < i \text{ and } f^{-1}(j) > f^{-1}(i) \right\} \right|,$$

i.e., v_i counts the number of elements of lesser value than i , but which appear to the right of i in the vector $f = [f_1, f_2, \dots, f_n]$. We note that $v_1 = 0$ always, and is thus redundant in the representation, but we keep it to make the notation simpler. From now on, we denote the factoradic representation of $f \in S_n$ by $\Phi(f) \in \mathbb{Z}_n!$, and the i th element of $\Phi(f)$ by $\Phi(f)_i$.

We now crucially observe that, in a systematic scheme, setting the levels of the first k cells determines exactly the first k entries of the factoradic representation of the permutation stored by the n cells. This is true regardless of the levels of the last $n-k$ cells. More succinctly, for any $f \in S_n$, and for all $1 \leq i \leq k \leq n$,

$$\Phi(f|^{[k]})_i = \Phi(f)_i.$$

Example 4. Let $n = 6$ and $k = 4$. Take

$$f = [6, 1, 3, 2, 5, 4] \in S_n.$$

We then have

$$\Phi(f) = [0, 0, 1, 0, 1, 5],$$

as well as

$$f|^{[k]} = [1, 3, 2, 4] \quad \text{and} \quad \Phi(f|^{[k]}) = [0, 0, 1, 0].$$

We observe that the first k coordinates of $\Phi(f)$ and $\Phi(f|^{[k]})$ are the same. \square

Another well-known fact (used by [1], [9]) is the following metric embedding:

$$d_K(f, g) \geq d_1(\Phi(f), \Phi(g)) = \sum_{i=1}^n |\Phi(f)_i - \Phi(g)_i|, \quad (1)$$

where d_K is Kendall's τ -distance, and d_1 is the ℓ_1 -distance. The following lemma gives a more refined version of (1), taking into account the partition into information symbols and redundancy symbols.

Lemma 5. Given $f, g \in S_n$, and $1 \leq k \leq n$,

$$d_K(f, g) \geq d_K(f|^{[k]}, g|^{[k]}) + \sum_{i=k+1}^n |\Phi(f)_i - \Phi(g)_i|.$$

Proof: The proof is by induction on $r = n - k$. As the base case, the inequality is clearly satisfied for $r = 0$, i.e., $n = k$. Now consider the inductive step. Suppose that the inequality holds for some $r - 1 = n - k - 1$, and we will now show that it also holds for $r = n - k$.

Consider a sequence of $d_K(f, g)$ adjacent transpositions that changes the permutation f into the permutation g . Of these transpositions, assume that α adjacent transpositions involve the integer n , and β adjacent transpositions do not involve n . Clearly,

$$d_K(f, g) = \alpha + \beta.$$

Since the integer n needs to be moved from position $n - \Phi(f)_n$ to position $n - \Phi(g)_n$, we get

$$\alpha \geq |\Phi(f)_n - \Phi(g)_n|.$$

Note that those adjacent transpositions that involve n do not change the relative order of the integers $[n - 1]$ in the permutation. Thus, to transform the integers $[n - 1]$ from their relative order in f to their relative order in g , by the induction assumption, we get

$$\beta \geq d_K(f|^{[k]}, g|^{[k]}) + \sum_{i=k+1}^{n-1} |\Phi(f)_i - \Phi(g)_i|.$$

That leads to the conclusion. \blacksquare

Example 6. Let $n = 3$ and $k = 2$ and consider

$$f = [1, 3, 2] \quad \text{and} \quad g = [2, 1, 3].$$

In this case, the inequality of Lemma 5 becomes an equality since

$$\begin{aligned} d_K([1, 3, 2], [2, 1, 3]) &= 2 \\ &= 1 + |1 - 0| = d_K([1, 2], [2, 1]) + |\Phi(f)_3 - \Phi(g)_3|. \end{aligned}$$

The equality, however, does not always hold. For instance, if

$$f' = [1, 3, 2] \quad \text{and} \quad g' = [2, 3, 1],$$

we get

$$\begin{aligned} d_K([1, 3, 2], [2, 3, 1]) &= 3 \\ &> 1 + |1 - 1| = d_K([1, 2], [2, 1]) + |\Phi(f)_3 - \Phi(g)_3|. \end{aligned}$$

\square

We now present an inequality for ball sizes in S_n , which will be useful for the analysis of systematic codes. Given a permutation $f \in S_n$, the ball of radius r centered at f , is defined by

$$\mathfrak{B}_r(f) = \{g \in S_n \mid d_K(f, g) \leq r\},$$

for any $0 \leq r \leq \binom{n}{2}$. We recall that the maximum distance for any two permutations in S_n is $\binom{n}{2}$ (for example, see [9]). A simple relabeling argument suffices to show that the size of a ball does not depend on the choice of its center. Therefore, we will use $|\mathfrak{B}_r(n)|$ to denote $|\mathfrak{B}_r(f)|$ for any $f \in S_n$.

An exact expression for $|\mathfrak{B}_r(n)|$ is known [14]. However, for our purposes, we will use the inequality of the following lemma.

Lemma 7. For all $n \geq 1$ and $0 \leq r \leq \binom{n}{2}$,

$$|\mathfrak{B}_r(n)| \leq \binom{n+r-1}{n-1}.$$

Proof: Since the center of a ball does not affect its size, consider the ball centered at the identity, $\mathfrak{B}_r(\text{Id})$. It follows from (1) that

$$|\mathfrak{B}_r(\text{Id})| \leq |\{f \in S_n \mid d_1(\Phi(f), \Phi(\text{Id})) \leq r\}|. \quad (2)$$

Since, conveniently, $\Phi(\text{Id})$ is the all-zero vector, we have for any $f \in S_n$ that

$$d_1(\Phi(f), \Phi(\text{Id})) = \sum_{i=1}^n \Phi(f)_i.$$

We further note that $\Phi(f)_1 = 0$ always.

Thus, the right-hand side of (2) is upper bounded by the number of non-negative-integer vectors of length $n - 1$ whose entry sum is at most r . This is easily seen to be the same as the number of ways r identical balls can be thrown into n non-identical bins, and hence,

$$|\mathfrak{B}_r(n)| \leq \binom{n+r-1}{n-1}.$$

\blacksquare

B. Systematic One-Error-Correcting Codes

We start by presenting two constructions for systematic $[k + 2, k, 3]$ codes, capable of correcting a single error. The first construction uses a direct manipulation of the permutations to construct the codewords, and is somewhat restricted in its choice of parameters. In contrast, the second construction uses a metric embedding technique, and applies to all parameters. We then show the codes are optimal unless perfect one-error-correcting codes exist.

Construction A. Let $k \geq 3$ be an integer such that either k or $k + 1$ is a prime. For any $f \in S_k$, and $j \geq 1$ an integer, we define the following function:

$$\rho_j(f) = \left(\sum_{i=1}^k (2i - 1)^j f(i) \right) \bmod m, \quad (3)$$

where $m = k$ if k is a prime and $m = k + 1$ if $k + 1$ is a prime.

We construct the code

$$C = \left\{ f \in S_{k+2} \mid \Phi(f)_{k+j} = \rho_j(f|^{[k]}) \text{ for all } j \in [2] \right\}.$$

\square

Theorem 8. The code C from Construction A is a systematic $[k + 2, k, 3]$ code in Kendall's τ -metric.

Proof: We easily observe that the information symbols $[k]$ are unconstrained, and so

$$\left\{ f|^{[k]} \mid f \in C \right\} = S_k.$$

Furthermore, since a choice of the order of the information symbols determines the position of the two redundancy symbols uniquely, we also have $|C| = k!$.

It now only remains to show that the minimum distance of C is 3. We know that either k is a prime, or $k + 1$ is a prime. Let us first handle the former case. Let $f, g \in C$ be two codewords, $f \neq g$. We divide our proof into three cases, depending on $d_K(f|^{[k]}, g|^{[k]})$.

a) *Case 1:* $d_K(f|^{[k]}, g|^{[k]}) = 1$. In this case, we can write

$$\begin{aligned} f|^{[k]} &= [a_1, a_2, \dots, a_i, a_{i+1}, \dots, a_k], \\ g|^{[k]} &= [a_1, a_2, \dots, a_{i+1}, a_i, \dots, a_k]. \end{aligned}$$

for some $i \in [k-1]$, i.e., $f|^{[k]}$ and $g|^{[k]}$ differ by an adjacent transposition of the i th and $(i+1)$ st elements.

Let us now define $\Delta = a_{i+1} - a_i$. It follows that

$$\Phi(f)_{k+1} - \Phi(g)_{k+1} \equiv 2\Delta \pmod{k}.$$

Since $1 \leq |\Delta| \leq k-1$ and $k \geq 3$ is a prime, we know that 2Δ is not a multiple of k . As a result, we get

$$|\Phi(f)_{k+1} - \Phi(g)_{k+1}| \geq 1.$$

Similarly, we have

$$\begin{aligned} \Phi(f)_{k+2} - \Phi(g)_{k+2} &\equiv (2i-1)^2 a_i + (2i+1)^2 (a_i + \Delta) \\ &\quad - (2i-1)^2 (a_i + \Delta) - (2i+1)^2 a_i \\ &\equiv 8i\Delta \pmod{k}. \end{aligned}$$

Again, $8i\Delta$ is not a multiple of k since $1 \leq i, |\Delta| \leq k-1$ and $k \geq 3$ is a prime. This implies that

$$|\Phi(f)_{k+2} - \Phi(g)_{k+2}| \geq 1.$$

Thus, by Lemma 5, we get

$$\begin{aligned} d_K(f, g) &\geq d_K(f|^{[k]}, g|^{[k]}) + \sum_{i=k+1}^{k+2} |\Phi(f)_i - \Phi(g)_i| \\ &\geq 1 + 1 + 1 = 3. \end{aligned}$$

b) *Case 2:* $d_K(f|^{[k]}, g|^{[k]}) = 2$. Let us denote

$$f|^{[k]} = [a_1, a_2, \dots, a_k].$$

By our assumption, there exist $1 \leq i, j \leq k-1$ such g is obtained from f as a result of two adjacent transpositions: one exchanging locations i and $i+1$, and one exchanging locations j and $j+1$. We distinguish between two cases.

In the first case, $\{i, i+1\} \cap \{j, j+1\} = \emptyset$. Without loss of generality, assume $i < j$, and then we have

$$g|^{[k]} = [a_1, \dots, a_{i+1}, a_i, \dots, a_{j+1}, a_j, \dots, a_k].$$

Let us define $\Delta_1 = a_{i+1} - a_i$, and $\Delta_2 = a_{j+1} - a_j$. Then we get

$$\Phi(f)_{k+1} - \Phi(g)_{k+1} \equiv 2(\Delta_1 + \Delta_2) \pmod{k}.$$

If $\Delta_1 + \Delta_2$ is not a multiple of k , then by the same reasoning as before,

$$|\Phi(f)_{k+1} - \Phi(g)_{k+1}| \geq 1.$$

If $\Delta_1 + \Delta_2$ is a multiple of k , then we can write $\Delta_2 \equiv -\Delta_1 \pmod{k}$. Hence,

$$\begin{aligned} \Phi(f)_{k+2} - \Phi(g)_{k+2} &\equiv (2i-1)^2 a_i + (2i+1)^2 (a_i + \Delta_1) \\ &\quad + (2j-1)^2 a_j + (2j+1)^2 (a_j - \Delta_1) \\ &\quad - (2i-1)^2 (a_i + \Delta_1) - (2i+1)^2 a_i \\ &\quad - (2j-1)^2 (a_j - \Delta_1) - (2j+1)^2 a_j \\ &\equiv 8(j-i)\Delta_1 \pmod{k}. \end{aligned}$$

Since $8(j-i)\Delta_1$ is not a multiple of k , we have

$$|\Phi(f)_{k+2} - \Phi(g)_{k+2}| \geq 1.$$

In the second case, $\{i, i+1\} \cap \{j, j+1\} \neq \emptyset$. Thus, either

$$g|^{[k]} = [a_1, \dots, a_{i+2}, a_i, a_{i+1}, \dots, a_k],$$

or

$$g|^{[k]} = [a_1, \dots, a_{i+1}, a_{i+2}, a_i, \dots, a_k],$$

for some $i \in [k-2]$. By defining $\Delta_1 = a_{i+2} - a_{i+1}$ and $\Delta_2 = a_{i+2} - a_i$ in the first case, or $\Delta_1 = a_{i+1} - a_i$ and $\Delta_2 = a_{i+2} - a_i$ in the second case, and with the same arguments as above, it can be proved that either

$$|\Phi(f)_{k+1} - \Phi(g)_{k+1}| \geq 1,$$

or

$$|\Phi(f)_{k+2} - \Phi(g)_{k+2}| \geq 1.$$

Combining all the cases together, by Lemma 5, we get

$$\begin{aligned} d_K(f, g) &\geq d_K(f|^{[k]}, g|^{[k]}) + \sum_{i=k+1}^{k+2} |\Phi(f)_i - \Phi(g)_i| \\ &\geq 2 + 1 = 3. \end{aligned}$$

c) *Case 3:* $d_K(f|^{[k]}, g|^{[k]}) \geq 3$. This is the easiest case, since by Lemma 5,

$$d_K(f, g) \geq d_K(f|^{[k]}, g|^{[k]}) \geq 3.$$

Finally, we note that if $k+1$ is a prime, we can repeat the proof in its entirety, replacing $\text{mod } k$ with $\text{mod } (k+1)$. ■

Before continuing to the next construction we would like to consider encoding and decoding algorithms for the code from Construction A. For the encoding procedure, we start by mapping an integer from $\mathbb{Z}_{k!}$ to a permutation $f' \in S_k$. This may be accomplished in linear time [17]. Then, using the description of Construction A, the two redundancy symbols are easily placed in their correct position, and we receive a codeword $f \in C$ such that $f|^{[k]} = f'$.

Decoding may be done efficiently as well. Assume $f \in C \subseteq S_{k+2}$ was transmitted, while $g \in S_{k+2}$ was received, where $d_K(f, g) \leq 1$. A trivial decoding algorithm can check the $k+2$ permutation in the ball of radius 1 centered around g , and find the unique codeword f in it. This algorithm takes $O(k^2)$ steps.

We can do better than that, using the decoding algorithm we now describe. Let $\hat{g} \in C$ be the unique codeword in C having the same order of information symbols as g , i.e., $\hat{g}|^{[k]} = g|^{[k]}$. If $d_K(\hat{g}, g) \leq 1$, then $f = \hat{g}$ is the correct decoding. Otherwise, $d_K(f|^{[k]}, g|^{[k]}) = 1$, and we can write

$$\begin{aligned} f|^{[k]} &= [a_1, \dots, a_i, a_{i+1}, \dots, a_k], \\ g|^{[k]} &= [a_1, \dots, a_{i+1}, a_i, \dots, a_k], \end{aligned}$$

for some $i \in [k-1]$.

Since a single adjacent transposition changed the order of two information symbols, we deduce no redundancy symbols were moved, and thus,

$$\Phi(f)_{k+1} = \Phi(g)_{k+1} \quad \text{and} \quad \Phi(f)_{k+2} = \Phi(g)_{k+2}.$$

According to the proof of Theorem 8,

$$\begin{aligned}\Phi(g)_{k+1} - \Phi(\hat{g})_{k+1} &\equiv 2(a_{i+1} - a_i) \pmod{m}, \\ \Phi(g)_{k+2} - \Phi(\hat{g})_{k+2} &\equiv 8i(a_{i+1} - a_i) \pmod{m},\end{aligned}$$

where m is the prime in $\{k, k+1\}$. Combining the two equations together we get

$$\Phi(g)_{k+2} - \Phi(\hat{g})_{k+2} \equiv 4i(\Phi(g)_{k+1} - \Phi(\hat{g})_{k+1}) \pmod{m},$$

and we can easily solve for i , thus recovering the coordinate of the adjacent transposition. This decoding algorithm runs in $O(k)$ steps. We illustrate the decoding algorithm with the following example.

Example 9. Let $k = 4$, and assume we would like to encode $[4, 1, 3, 2]$. Thus, by Construction A, we look for a permutation $f \in S_6$ such that

$$\begin{aligned}\Phi(f)_5 &= \left(\sum_{i=1}^4 (2i-1)f(i) \right) \pmod{5} = 1, \\ \Phi(f)_6 &= \left(\sum_{i=1}^4 (2i-1)^2 f(i) \right) \pmod{5} = 1.\end{aligned}$$

We therefore transmit

$$f = [4, 1, 3, 5, 6, 2],$$

and let us assume the received permutation is

$$g = [4, 3, 1, 5, 6, 2],$$

due to an adjacent transposition in positions 2 and 3. We extract the information symbols from g to obtain,

$$g|^{[k]} = [4, 3, 1, 2],$$

and use that to construct the codeword

$$\hat{g} = [4, 6, 3, 5, 1, 2].$$

Since $d_K(\hat{g}, g) > 1$, we deduce that two information symbols changed positions. Since

$$\begin{aligned}\Phi(g)_{k+1} &= 1 & \Phi(g)_{k+2} &= 1 \\ \Phi(\hat{g})_{k+1} &= 2 & \Phi(\hat{g})_{k+2} &= 4,\end{aligned}$$

we solve

$$1 - 4 \equiv 4i(1 - 2) \pmod{5},$$

resulting in the correct positions of the adjacent transposition, $i = 2$ and $i + 1 = 3$. \square

Another strategy for constructing rank-modulation codes for Kendall's τ -metric, which was already employed by [1], [9], is to first construct a code C^* with minimum ℓ_1 -distance d in \mathbb{Z}^n , and then take

$$C = \Phi^{-1}(C^* \cap \mathbb{Z}_n!),$$

i.e., exactly those permutations whose factoradic representations are $C^* \cap \mathbb{Z}_n!$. Since by (1), the distance can only increase, the resulting set of permutations is a code with minimum Kendall's τ -distance of at least d . The main challenge with this approach is to ensure a large intersection of C^* with $\mathbb{Z}_n!$.

For the construction of systematic codes we shall employ the same methods, however, now we have an additional challenge: we also require the intersection $C^* \cap \mathbb{Z}_n!$ to have at least one vector for each possible prefix from $\mathbb{Z}_k!$.

Construction B. Let $k \geq 2$ be some integer. For a vector $x = (x_1, x_2, \dots, x_{k+1}) \in \mathbb{Z}^{k+1}$ we denote

$$s_m(x) = \left(\sum_{i=1}^m ix_i \right) \pmod{(2k+3)}.$$

We construct a subset $C' \subseteq \mathbb{Z}^{k+1}$ defined by

$$\begin{aligned}C' &= \{x \in \mathbb{Z}^{k+1} \mid x_k = \lfloor s_{k-1}(2x)/3 \rfloor, \\ &\quad x_{k+1} = s_{k-1}(2x) \pmod{3}\}.\end{aligned}$$

We denote by

$$C^* = \{(0, x_1, x_2, \dots, x_{k+1}) \mid (x_1, x_2, \dots, x_{k+1}) \in C'\},$$

the prepending of 0 to all the codewords of C' . The constructed code is

$$C = \Phi^{-1}(C^* \cap \mathbb{Z}_{k+2}!).$$

\square

Theorem 10. For all $k \geq 2$, the code C from Construction B is a $[k+2, k, 3]$ systematic code in Kendall's τ -metric.

Proof: Consider the perfect $(k+1)$ -dimensional single-error-correcting code in the ℓ_1 -metric described by Golomb and Welch in [7] and given by,

$$C'' = \left\{ x = (x_1, x_2, \dots, x_{k+1}) \in \mathbb{Z}^{k+1} \mid s_{k+1}(x) = 0 \right\}.$$

We contend that $C' \subseteq C''$, i.e., that C' is also a single-error-correcting code in the ℓ_1 -metric. Indeed, let $x = (x_1, \dots, x_{k+1}) \in C'$ be a codeword in C' . Then, noting that

$$k \equiv 3(k+1) \pmod{2k+3},$$

and working modulo $2k+3$, and we get

$$\begin{aligned}s_{k+1}(x) &\equiv s_{k-1}(x) + kx_k + (k+1)x_{k+1} \\ &\equiv s_{k-1}(x) + k \lfloor s_{k-1}(2x)/3 \rfloor \\ &\quad + (k+1)(s_{k-1}(2x) \pmod{3}) \\ &\equiv s_{k-1}(x) + (k+1)(3 \lfloor s_{k-1}(2x)/3 \rfloor \\ &\quad + (s_{k-1}(2x) \pmod{3})) \\ &\equiv s_{k-1}(x) + (k+1)s_{k-1}(2x) \\ &\equiv s_{k-1}(x) + 2(k+1)s_{k-1}(x) \\ &\equiv (2k+3)s_{k-1}(x) \equiv 0 \pmod{2k+3}.\end{aligned}$$

Thus, $x \in C''$, and so $C' \subseteq C''$.

We note how the first $k-1$ coordinates of the codewords of C' are unconstrained. Thus, for all $1 \leq i \leq k-1$ we can set $x_i \in \mathbb{Z}_{i+1}$ arbitrarily in any one of $k!$ ways. Furthermore, for any $x \in C'$,

$$0 \leq \lfloor s_{k-1}(2x)/3 \rfloor \leq \frac{2(k+1)}{3} \leq k,$$

as well as

$$0 \leq s_{k-1}(2x) \pmod{3} \leq k+1.$$

It follows that $x_k \in \mathbb{Z}_{k+1}$ and $x_{k+1} \in \mathbb{Z}_{k+2}$. Hence, after prepending a 0 to the codewords to obtain C^* , we get

$$|C^* \cap \mathbb{Z}_{k+2}| = k!.$$

Finally, prepending the 0 does not change the minimum distance, and so C^* has minimum ℓ_1 -distance of 3, and therefore, so does the final constructed code $C = \Phi^{-1}(C^* \cap \mathbb{Z}_{k+2})$. ■

Encoding the code from Construction B is extremely easy. In the factoradic representation we arbitrarily fill in the first $k-1$ entries. The last two digits are determined by the first $k-1$ digits, and a 0 is then prepended. We then convert the factoradic representation to a permutation, which is the desired codeword. The entire procedure takes $O(k)$ steps if we use [17] to convert from the factoradic representation to permutations.

The decoding process is simple as well. Given a permutation read from the channel, we first translate it to its factoradic representation and remove the leading 0. The remaining $k+1$ coordinates are decoded using any simple procedure for decoding the Golomb-Welch code [7]. Again, the entire procedure takes $O(n)$ steps.

We note that the two constructions are not equivalent. As an example, the $[5, 3, 3]$ code from Construction A contains the codewords

$$f = [1, 4, 3, 2, 5] \quad \text{and} \quad g = [2, 3, 4, 1, 5].$$

However,

$$\Phi(f) = [0, 0, 1, 2, 0] \quad \text{and} \quad \Phi(g) = [0, 1, 1, 1, 0].$$

Since

$$d_1(\Phi(f), \Phi(g)) = 2,$$

the code cannot have originated from Construction B.

An obvious question to ask is how good are the parameters of the codes presented in Construction A and Construction B. Any (n, M, d) code (systematic or not) has to satisfy the ball-packing bound:

$$M \leq \frac{n!}{|\mathcal{B}_r(n)|}, \quad (4)$$

where $r = \lfloor (d-1)/2 \rfloor$. Codes attaining (4) with equality are called *perfect*. We thus reach the following simple corollary:

Corollary 11. *The $[k+2, k, 3]$ systematic codes of Construction A and Construction B have optimal size, unless perfect systematic one-error-correcting codes exist in Kendall's τ -metric.*

Example of perfect codes in other metrics are quite rare (see [16]). In Kendall's τ -metric there is a simple $(3, 2, 3)$ that is perfect:

$$C = \{[1, 2, 3], [3, 2, 1]\}.$$

This code is also systematic, i.e., a $[3, 2, 3]$ -code. However, beside this code, no other perfect code has been found yet. It was recently shown in [3], that no perfect codes exist in S_n under Kendall's τ -metric when n is a prime, or when $4 \leq n \leq 10$.

To summarize, the $[k+2, k, 3]$ codes presented have minimal redundancy among systematic codes, unless there exists

a perfect systematic $[k+1, k, 3]$ one-error-correcting code. Furthermore, compared with the one-error-correcting code presented in [9], the codes presented here have more efficient encoding and decoding algorithms.

C. Multi-Error-Correcting Codes

After studying systematic one-error-correcting codes, we turn to consider systematic codes capable of correct more than one error. We will first describe an explicit construction for a wide range of parameters, and then turn to a greedy algorithm leading us to prove a non-constructive existence result.

The systematic one-error-correcting code in Construction A may be generalized in a straightforward way: for $1 \leq k \leq n$ and $r \geq 1$ integers we define,

$$C = \left\{ f \in S_{k+r} \mid \Phi(f)_{k+j} = \rho_j(f^{[k]}) \text{ for all } j \in [r] \right\},$$

where $\rho_j(\cdot)$ is given by (3). This gives us a family of codes, including a $[10, 4, 5]$ systematic code, and a $[14, 4, 7]$ systematic code. However, a general analysis of these codes is difficult.

We therefore return to the strategy of metric embedding: an ℓ_1 -metric code is constructed in such a way as to allow all possible values in the first few information entries, and then the other positions are determined as a function of the information entries.

Construction C. *Let p be a prime, $m \geq 2$ an integer, $1 \leq t \leq \frac{p-3}{2}$ also an integer, and*

$$\max(p^{m-1}, p + tm - 1) \leq n \leq p^m - 1.$$

Arbitrarily choose $\alpha_1, \alpha_2, \dots, \alpha_n$ to be n distinct non-zero elements of $\text{GF}(p^m)$. We define

$$H = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \vdots & \vdots & \dots & \vdots \\ \alpha_1^t & \alpha_2^t & \dots & \alpha_n^t \end{bmatrix}.$$

Viewing $\text{GF}(p^m)$ as the vector space $\text{GF}(p)^m$, we can think of any entry of the form α_i^j in H as a column vector of length m over $\text{GF}(p)$. Thus, we shall consider H to be a $(t+1)m \times n$ matrix over $\text{GF}(p)$. We denote

$$k = n - \text{rank } H.$$

We construct a subset $C' \subseteq \mathbb{Z}^n$ defined by

$$C' = \{x \in \text{GF}(p)^n \mid Hx \equiv 0 \pmod{p}\},$$

where the entries of Hx are computed modulo p .

We define the mapping $\mu : \mathbb{Z}^n \rightarrow \mathbb{Z}^{n+1}$ as follows,

$$\begin{aligned} \mu(x_1, x_2, \dots, x_n) &= \\ &= (0, x_1, \dots, x_k, x_{k+1} \bmod p, \dots, x_n \bmod p), \end{aligned}$$

i.e., prepending a zero and reducing the last $n-k$ entries modulo p . We then set

$$C^* = \{\mu(x) \mid x \in C'\}.$$

The constructed code is

$$C = \Phi^{-1}(C^* \cap \mathbb{Z}_{n+1}!).$$

□

Theorem 12. *The code C from Construction C is an $[n+1, k+1, 2t+2]$ systematic code in Kendall's τ -metric capable of correcting t errors. Furthermore, the code's redundancy satisfies $n-k \leq tm+1$.*

Proof: The matrix H is nothing but the parity-check matrix for a BCH code over $\text{GF}(p)$. Since the code is linear, we can find a $k \times n$ generator matrix G for the code, and in particular, we can require that it be systematic, i.e.,

$$G = [I_k | A],$$

where I_k is the $k \times k$ identity matrix, and A is a $k \times (n-k)$ matrix over $\text{GF}(p)$. As a side note, getting to this systematic form, we may be required to permute the coordinates of the code. Since the order of elements $\alpha_1, \dots, \alpha_n$, which are used to construct H , is irrelevant, we assume it is chosen so that no change of order of coordinates is required.

Let us denote by C'' the code whose generator matrix is G . We recall the definition of the Lee-distance measure over $\text{GF}(p)$: given two vectors $x, x' \in \text{GF}(p)^n$,

$$d_L(x, x') = \sum_{i=1}^n \min(|x_i - x'_i|, |x'_i - x_i|),$$

where subtraction is done in $\text{GF}(p)$. It was shown in [19], that the minimum Lee distance of the code C'' is at least $2t+2$.

Our next goal is to transform C'' to a code over \mathbb{Z}^n with a minimum ℓ_1 -distance guarantee. Since we have the code C'' reside in the n -dimensional cube $\text{GF}(p)^n$, we can place copies of that cube and tile the entire space \mathbb{Z}^n . This is known as Construction A of [15], and the resulting code is exactly

$$C' = \{x \in \text{GF}(p)^n \mid Hx \equiv 0 \pmod{p}\},$$

where the entries of Hx are computed modulo p . Again, by [15], the codewords of C' are spanned (using linear combinations with integer coefficients) by the generating matrix

$$G' = \begin{bmatrix} I_k & A \\ 0 & pI_{n-k} \end{bmatrix}.$$

Thus, the minimum Lee distance of $2t+2$ between codewords of C'' guarantees a minimum ℓ_1 -distance of $2t+2$ between codewords of C' .

A quick inspection of G' reveals that, due to the first k rows, any prefix of k integers may be completed to a length- n codeword in C' . Furthermore, given a codeword in C' , by reducing its last $n-k$ entries modulo p we obtain another codeword of C' , due to the last $n-k$ rows of G' . It follows that

$$C^* = \{\mu(x) \mid x \in C'\},$$

is a subset of the codewords of C' with a 0 prepended, and that

$$|C^* \cap \mathbb{Z}_{n+1}!| = (k+1)!.$$

Hence, C is indeed an $[n+1, k+1, 2t+2]$ systematic code in Kendall's τ -metric.

Finally, it is well known (see also [19]) that when H is viewed as a $(t+1)m \times n$ matrix over $\text{GF}(p)$,

$$n-k = \text{rank } H \leq tm+1.$$

■

Again, encoding and decoding are easily done. For an encoding procedure, take any vector $(0|u) \in \mathbb{Z}_{k+1}!$ and map it to

$$(0|u) \mapsto (0|u|uA \bmod p) \in C^*.$$

The permutation whose factoradic representation is given by this vector is the encoded permutation.

For a decoding procedure, map the received permutation to its factoradic representation, and use the decoding for the Lee-metric code (essentially, a BCH decoding procedure) given in [19].

We also note that for the least redundancy, we would like to choose $m=2$ in Construction C. Using Bertrand's postulate, that any interval $[s, 2s]$ contains a prime, s a positive integer, we can show the following corollary:

Corollary 13. *For any $t \geq 1$, and $n \geq 6t+5$, there exists a prime p , such that the requirements of Construction C are satisfied with $m=2$, and therefore there exists an $[n+1, k+1, 2t+2]$ systematic code with redundancy at most $2t+1$.*

Along the same lines, but using two embeddings, one after the other, we present a construction transforming systematic binary codes under the Hamming metric, into systematic codes of permutations under Kendall's τ -metric. The construction is a simple modification of the construction given in [18].

The main idea for the first embedding is to use a mapping $\mathcal{G}_m : \mathbb{Z}_{2^m} \rightarrow \{0,1\}^m$ such that for any two integers $t_1, t_2 \in \mathbb{Z}_{2^m}$,

$$|t_1 - t_2| \geq d_H(\mathcal{G}_m(t_1), \mathcal{G}_m(t_2)), \quad (5)$$

where $d_H(\cdot, \cdot)$ denotes the Hamming distance function. By convention, \mathcal{G}_0 is the mapping returning the unique vector of length 0. A simple way of creating such a mapping is to use the encoding function for an optimal Gray code (see [20] for a survey of Gray codes).

Construction D. *Let C' be an $(n', 2^{k'}, d)$ binary systematic code in the Hamming metric, where the first k' coordinates are systematic. Furthermore, let k and n be integers such that*

$$k' = \sum_{i=1}^k \lceil \log_2 i \rceil,$$

$$n' = \sum_{i=1}^k \lceil \log_2 i \rceil + \sum_{i=k+1}^n \lfloor \log_2 i \rfloor.$$

We conveniently define

$$\lambda(i) = \begin{cases} \lceil \log_2 i \rceil & 1 \leq i \leq k, \\ \lfloor \log_2 i \rfloor & k+1 \leq i \leq n. \end{cases}$$

We now construct the following code,

$$C = \left\{ f \in S_n \mid \mathcal{G}_{\lambda(1)}(\Phi(f)_1) \parallel \dots \parallel \mathcal{G}_{\lambda(n)}(\Phi(f)_n) \in C' \right\}.$$

where \parallel denotes vector concatenation. In particular, the notation implies that for any $f \in C$,

$$0 \leq \Phi(f)_i \leq 2^{\lfloor \log_2 i \rfloor} - 1 \leq i - 1,$$

for all $k + 1 \leq i \leq n$. \square

Theorem 14. *The code C from Construction D is an $[n, k, d]$ systematic code in Kendall's τ -metric.*

Proof: The length of the code is obviously n . Let us try to build a codeword $f \in C$. We note that the first k symbols of $\Phi(f)$ form a binary vector of length k' after being Gray-mapped and concatenated. Since the first k' bits of the code C' are systematic, any such k' -prefix may be uniquely completed to form a codeword in C' by adding appropriate $n' - k'$ redundancy bits. These redundancy bits can be divided into sets of size $\lfloor \log_2 i \rfloor$, with $k + 1 \leq i \leq n$. Thus, the reverse Gray mapping of these sets uniquely determines $\Phi(f)_{k+1}, \dots, \Phi(f)_n$, and therefore, f as well. It follows that C is indeed a systematic code of length n and k information symbols.

Finally, let $f, g \in C$ be two distinct codewords. Then, using (1) and (5) we get

$$\begin{aligned} d_K(f, g) &\geq \sum_{i=1}^n |\Phi(f)_i - \Phi(g)_i| \\ &\geq \sum_{i=1}^n d_H(\mathcal{G}_{\lambda(i)}(\Phi(f)_i), \mathcal{G}_{\lambda(i)}(\Phi(g)_i)) \\ &\geq d. \end{aligned}$$

Thus, C is an $[n, k, d]$ systematic code. \blacksquare

We turn to present a generic scheme for constructing an $[n, k, d]$ systematic codes. The scheme is a simple adaptation of the Gilbert-Varshamov lower bound. Although it is beyond the scope of this paper to obtain efficient encoding and decoding algorithms for it, the analysis of this scheme is very useful for proving the existence of codes with certain parameters, and for deriving the capacity of systematic codes.

Construction E. *Let $2 \leq k < n$ and $d \geq 1$ be integers. We define $C_0 = \emptyset$. For all $i \geq 1$, the algorithm searches for $f \in S_n$ such that*

$$\min_{g \in C_{i-1}} d_K(f, g) \geq d, \quad (6)$$

and

$$f^{[k]} \notin \{g^{[k]} \mid g \in C_{i-1}\}. \quad (7)$$

If we can continue the process, increasing i by 1 at each iteration, and reach $i = k!$, then the constructed code is $C = C_{k!}$. Otherwise, we declare failure. \square

Theorem 15. *If the algorithm of Construction E succeeds, the resulting code C is a systematic $[n, k, d]$ code in Kendall's τ -metric.*

Proof: The proof is straightforward. We start with an empty set, and at each stage we add a codeword that is not too close to previously chosen codewords. Thus, the minimum distance of the resulting code is d . The second requirement at

each step, is that the information symbols do not repeat those of a previously chosen codeword. Thus,

$$\{f^{[k]} \mid f \in C\} = S_k,$$

and the code is systematic. \blacksquare

For Construction E to succeed, the number of redundancy symbols, $n - k$, needs to be sufficiently large. In the following theorem, we derive a bound for these parameters.

Theorem 16. *Construction E can successfully build an $[n, k, d]$ systematic code if*

$$\sum_{i=1}^{d-1} \binom{k+i-2}{i} \binom{d-i-1+n-k}{n-k} 2^{\min(d-i-1, n-k)} < \frac{n!}{k!}.$$

Proof: For any permutations $h \in S_k$ there are exactly $n!/k!$ permutations $f \in S_n$ such that $f^{[k]} = h$. At each step i of the algorithm we shall arbitrarily choose $h \in S_k$ such that

$$h \notin \{g^{[k]} \mid g \in C_{i-1}\}.$$

We shall then try to find $f \in S_n$ such that $f^{[k]} = h$, i.e., f satisfies requirement (7). Our goal is to show there is at least one such f that also satisfies the requirement of (6).

Given any such $h \in S_k$, let us upper bound the number of permutations f such that $f^{[k]} = h$ but f does not satisfy (6). Let $g \in C_{i-1}$ be a codeword chosen in some previous iteration, and assume $d_K(f, g) \leq d - 1$. Let us denote

$$d_K(f^{[k]}, g^{[k]}) = j \leq d - 1.$$

By Lemma 5, in order for us to have $d_K(f, g) \leq d - 1$, we must have

$$\sum_{t=1}^{n-k} |\Phi(f)_{k+t} - \Phi(g)_{k+t}| \leq d - j - 1.$$

Thus, we would like to count the number of integer vectors of length $n - k$, whose ℓ_1 weight is at most $d - j - 1$. Choosing the magnitudes of the entries of such a vector is equivalent to the number of ways $d - j - 1$ identical balls can be placed in $n - k + 1$ non-identical bins. We also need to choose the sign for the non-zero entries of such a vector, and there are at most $\min(d - j - 1, n - k)$ such entries. It follows, that an upper bound on the number of permutations $f \in S_n$ such that $f^{[k]} = h$, and $d_K(f^{[k]}, g^{[k]}) = j$ for the given g , is

$$\binom{d-j-1+n-k}{n-k} 2^{\min(d-j-1, n-k)}.$$

Let N_j denote the number of permutations $g \in C_{i-1}$ such that $d_K(f^{[k]}, g^{[k]}) = j$. If we had this number, then by a simple union bound, the total number of permutations $f \in S_n$ such that $f^{[k]} = h$, but (6) does not hold, is upper bounded by

$$\sum_{j=1}^{d-1} N_j \binom{d-j-1+n-k}{n-k} 2^{\min(d-j-1, n-k)}.$$

To continue our upper bound, we replace N_j with the larger N'_j , where N'_j denotes the number permutations $h' \in S_k$ such

that $d_K(h', h) = j$. Our upper bound is now

$$\sum_{j=1}^{d-1} N'_j \binom{d-j-1+n-k}{n-k} 2^{\min(d-j-1, n-k)}. \quad (8)$$

According to Lemma 7,

$$\sum_{t=0}^j N'_t = |\mathfrak{B}_j(k)| \leq \binom{k+j-1}{k-1}.$$

In this case, it is not hard to prove that (8) is maximized when

$$N'_j = \binom{k+j-1}{k-1} - \binom{k+j-2}{k-1} = \binom{k+j-2}{k-2},$$

for $k \geq 2$ and $1 \leq j \leq d-1$, since

$$F(j) = 2^{\min(d-j-1, n-k)} \binom{d-j-1+n-k}{n-k},$$

is a decreasing function in j .

As a result, the number of permutations $f \in S_n$, such that $f|^{[k]} = h$, but (6) does not hold, is upper bounded by

$$\sum_{i=1}^{d-1} \binom{k+i-2}{i} \binom{d-i-1+n-k}{n-k} 2^{\min(d-i-1, n-k)}. \quad (9)$$

Since the total number of permutations $f \in S_n$ such that $f|^{[k]} = h$ is $n!/k!$, if (9) is strictly less than $n!/k!$ then there exists a permutation f satisfying (6). Since we did not restrict h in any way, this conclusion holds for any iteration of the algorithm, and the algorithm succeeds. ■

Example 17. When $d = 3$ and $n = k + 2$, the inequality of Theorem 16 can be simplified to

$$6 \binom{k-1}{1} + \binom{k}{2} < (k+1)(k+2),$$

which holds for any $k \geq 2$. Therefore, there exists a $[k+2, k, 3]$ systematic code for any $k \geq 2$. Note that this result is consistent with the codes built in Construction A and Construction B. □

Example 18. When $d = 4$ and $n = k + 3$, the inequality of Theorem 16 can be simplified to

$$40 \binom{k-1}{1} + 8 \binom{k}{2} + \binom{k+1}{3} < (k+1)(k+2)(k+3),$$

which holds for all $k \geq 2$. Therefore, there exists a $[k+3, k, 4]$ systematic code for any $k \geq 2$. □

We now prove our main existential result, which is non-constructive in nature, since it builds upon Construction E.

Theorem 19. There exists a $[k+d, k, d]$ systematic code in Kendall's τ -metric, for any $k \geq 2$ and $d \geq 1$.

Proof: Based on Theorem 16, to show that there exists a $[k+d, k, d]$ systematic code, we only need to prove

$$\sum_{i=1}^{d-1} \binom{k+i-2}{i} \binom{2d-1-i}{d} 2^{d-i-1} < \frac{(k+d)!}{k!}$$

for $k \geq 2$ and $d \geq 1$. We note that the case $d = 1$ is trivial, and so we will assume throughout the rest of the proof that

$d \geq 2$. Furthermore, to simplify the proof, we will prove a stronger claim,

$$\sum_{i=1}^{d-1} \binom{k+i}{i} \binom{2d-1-i}{d} 2^{d-i-1} < \frac{(k+d)!}{k!}. \quad (10)$$

Let us define

$$\psi_d(k) = \frac{k!}{(k+d)!} \sum_{i=1}^{d-1} \binom{k+i}{i} \binom{2d-1-i}{d} 2^{d-i-1}.$$

We contend the $\psi_d(k)$ is non-increasing in k , and to prove this claim we consider $\psi_d(k+1)/\psi_d(k)$ and note that

$$\begin{aligned} \frac{\psi_d(k+1)}{\psi_d(k)} &= \frac{\binom{k}{d}}{\binom{k+1}{d}} \cdot \frac{\sum_{i=1}^{d-1} \binom{k+1+i}{i} \binom{2d-1-i}{d} 2^{d-i-1}}{\sum_{i=1}^{d-1} \binom{k+i}{i} \binom{2d-1-i}{d} 2^{d-i-1}} \\ &= \frac{\binom{k}{d}}{\binom{k+1}{d}} \cdot \frac{\sum_{i=1}^{d-1} \frac{k+1+i}{k+1} \binom{k+i}{i} \binom{2d-1-i}{d} 2^{d-i-1}}{\sum_{i=1}^{d-1} \binom{k+i}{i} \binom{2d-1-i}{d} 2^{d-i-1}} \\ &\leq \frac{k+1}{k+d+1} \cdot \frac{k+d}{k+1} < 1. \end{aligned}$$

Thus, $\psi_d(k)$ is indeed a non-increasing function of k . If $\psi_d(2) < 1$ for all $d \geq 2$, then for any $k, d \geq 2$, we surely have $\psi_d(k) < 1$, which proves (10). So our task is to prove $\psi_d(2) < 1$, namely,

$$\sum_{i=1}^{d-1} \binom{2+i}{i} \binom{2d-1-i}{d} 2^{d-i-1} < \frac{(2+d)!}{2!}, \quad (11)$$

for all $d \geq 2$.

For all $2 \leq d \leq 16$ we can show that the inequality holds by computing the exact values. In what follows, we show that the inequality also holds when $d > 16$. The left-hand side of (11) may be upper bounded by

$$\begin{aligned} \sum_{i=1}^{d-1} \binom{2+i}{i} \binom{2d-1-i}{d} 2^{d-i-1} &\leq \\ &\leq \frac{d(d+1)(d+2)}{2} \binom{2d-2}{d} 2^{d-2}. \end{aligned}$$

Thus, to prove (11), it suffices to prove

$$\binom{2d-2}{d} 2^{d-2} < (d-1)!.$$

We define

$$\tilde{\zeta}(d) = \frac{1}{(d-1)!} \binom{2d-2}{d} 2^{d-2}.$$

We can numerically check that $\tilde{\zeta}(17) < 1$, and since

$$\frac{\tilde{\zeta}(d+1)}{\tilde{\zeta}(d)} = \frac{4d(2d+1)}{d(d+1)(d-1)} < 1$$

for all $d > 16$, we have $\tilde{\zeta}(d) < 1$ for all $d > 16$, and this completes the proof. ■

D. Capacity of Systematic Codes

In this section, we prove that for rank modulation under Kendall's τ -metric, systematic error-correcting codes achieve the same capacity as general error-correcting codes.

In [1], Barg and Mazumdar derived the capacity of general error-correcting codes for rank modulation under Kendall's τ -metric. Let $A(n, d)$ denote the maximum size of an (n, M, d) code. We define the capacity of error-correcting codes of minimum distance d as

$$\text{cap}(d) = \lim_{n \rightarrow \infty} \frac{\ln A(n, d)}{\ln n!}.$$

It was shown in [1] that

$$\text{cap}(d) = \begin{cases} 1 & \text{if } d = O(n), \\ 1 - \epsilon & \text{if } d = \Theta(n^{1+\epsilon}) \text{ with } 0 < \epsilon < 1, \\ 0 & \text{if } d = \Theta(n^2). \end{cases}$$

Turning to systematic codes, let $k(n, d)$ denote the maximum number of information symbols in systematic codes of length n and minimum distance d . Such codes are $[n, k(n, d), d]$ systematic codes, and have $k(n, d)!$ codewords. The *capacity* of systematic codes of minimum distance d is defined as

$$\text{cap}_{\text{sys}}(d) = \lim_{n \rightarrow \infty} \frac{\ln k(n, d)!}{\ln n!}.$$

The following theorem shows that systematic codes have the same capacity as general codes.

Theorem 20. *The capacity of systematic codes of minimum distance d is*

$$\text{cap}_{\text{sys}}(d) = \begin{cases} 1 & \text{if } d = O(n), \\ 1 - \epsilon & \text{if } d = \Theta(n^{1+\epsilon}) \text{ with } 0 < \epsilon < 1, \\ 0 & \text{if } d = \Theta(n^2). \end{cases}$$

Proof: Since systematic codes are a special case of general error-correcting codes, we naturally have

$$\text{cap}_{\text{sys}}(d) \leq \text{cap}(d).$$

Thus, to prove the claim, all that remains is to prove the other direction of the inequality.

According to Theorem 16, there exists an $[n, k, d]$ systematic code if k is the maximum integer that satisfies

$$d \binom{k+d}{d} \binom{d+n-k}{n-k} 2^n < \frac{n!}{k!}. \quad (12)$$

That is because

$$\begin{aligned} \sum_{i=1}^{d-1} \binom{k+i-2}{i} \binom{d-i-1+n-k}{n-k} 2^{\min(d-i-1, n-k)} &\leq \\ &\leq d \binom{k+d}{d} \binom{d+n-k}{n-k} 2^n, \end{aligned}$$

for all $n > k \geq 2$ and $d \geq 2$.

For such k , we have $k(n, d) \geq k$. For convenience, let us assume

$$\alpha = \lim_{n \rightarrow \infty} \frac{k}{n},$$

is a constant. We also recall the well-known Stirling's approximation,

$$\ln(m!) = m \ln m - O(m).$$

Thus, if $\alpha > 0$,

$$\begin{aligned} \text{cap}_{\text{sys}}(d) &= \lim_{n \rightarrow \infty} \frac{\ln k(n, d)!}{\ln n!} \geq \lim_{n \rightarrow \infty} \frac{\ln k!}{\ln n!} \\ &= \lim_{n \rightarrow \infty} \frac{\alpha n \ln(\alpha n) - O(n)}{n \ln n - O(n)} = \alpha. \end{aligned}$$

To prove the final conclusion, we will show that

$$\alpha \geq \begin{cases} 1 & \text{if } d = O(n), \\ 1 - \epsilon & \text{if } d = \Theta(n^{1+\epsilon}), \\ 0 & \text{if } d = \Theta(n^2). \end{cases} \quad (13)$$

We note that the last case is trivial, and so we only have to prove the first two.

By our choice of k and (12), we have

$$d \binom{k+d+1}{d} \binom{d+n-k-1}{n-k-1} 2^n \geq \frac{n!}{(k+1)!}.$$

It follows that

$$\lim_{n \rightarrow \infty} \frac{\ln \left(d \binom{k+d+1}{d} \binom{d+n-k-1}{n-k-1} 2^n \right)}{\ln \left(\frac{n!}{(k+1)!} \right)} \geq 1. \quad (14)$$

To prove the first case of (13) assume $d = O(n)$. Again, by Stirling's approximation, (14) becomes,

$$\begin{aligned} 1 &\leq \lim_{n \rightarrow \infty} \frac{\ln \left(d \binom{k+d+1}{d} \binom{d+n-k-1}{n-k-1} 2^n \right)}{\ln \left(\frac{n!}{(k+1)!} \right)} \\ &= \lim_{n \rightarrow \infty} \frac{O(n)}{n \ln n - \alpha n \ln(\alpha n) - O(n)}. \end{aligned}$$

Since α is a constant, we must therefore have $\alpha = 1$.

For the second case, assume $d = \Theta(n^{1+\epsilon})$ for $0 < \epsilon < 1$. By applying Stirling's approximation to (14), after some tedious rearranging, we get

$$\begin{aligned} 1 &\leq \lim_{n \rightarrow \infty} \frac{\ln \left(d \binom{k+d+1}{d} \binom{d+n-k-1}{n-k-1} 2^n \right)}{\ln \left(\frac{n!}{(k+1)!} \right)} \\ &= \lim_{n \rightarrow \infty} \frac{\epsilon n \ln n - O(n)}{(1 - \alpha) n \ln n - O(n)}. \end{aligned}$$

Thus, $\alpha \geq 1 - \epsilon$, as we wanted to show. \blacksquare

IV. SYSTEMATIC CODES IN THE ℓ_∞ -METRIC

We recall that the definition of systematic codes in the ℓ_∞ -metric differs from that in Kendall's τ -metric. Intuitively, in an $[n, k, d]$ systematic code in the ℓ_∞ -metric, when taking the first k coordinates of the $k!$ codewords and relabeling the surviving k elements to $[k]$, we obtain every permutation of S_k exactly once.

The exact capacity for codes in the ℓ_∞ -metric is unknown. There is a large gap between the lower and upper bounds on the size of optimal codes, mainly due to the lack of an asymptotic expression for the size of balls in this metric. Thus,

to evaluate the parameters of our construction we will compare the rate of the constructed systematic codes with that of known general codes. Given an (n, M, d) code C in the ℓ_∞ -metric, its *rate* is defined as (see [23])

$$R(C) = \frac{\log_2 M}{n}.$$

Note that this definition is somewhat different than that for Kendall's τ -metric (see [1]).

We present two constructions for systematic codes, where the first is adequate for distances $d = O(1)$, and where the second is intended for the $d = \Theta(n)$ case.

Construction F. Let $1 \leq d \leq n$ be positive integers, and let $1 \leq k \leq \lceil n/d \rceil$ be an integer as well. We denote

$$A_{k,d} = \{1, 1+d, 1+2d, \dots, 1+(k-1)d\}.$$

We construct the code

$$C = \{(f_1, \dots, f_n) \in S_n \mid \{f_1, \dots, f_k\} = A_{k,d}, \\ f_{k+1} < f_{k+2} < \dots < f_n\}$$

□

Theorem 21. The code C from Construction F is an $[n, k, d]$ systematic code in the ℓ_∞ -metric.

Proof: It is immediately evident that $|C| = k!$. Furthermore, every two distinct codewords in C disagree on at least one of the first k coordinates. Since all entries in the first k coordinates leave a residue of 1 modulo d , the ℓ_∞ -distance between distinct codewords is at least d . Finally, the projection onto the first k coordinates is easily seen to provide all possible permutations from S_k exactly once. ■

The optimal choice of k in Construction F is obviously $k = \lceil n/d \rceil$, and it provides a code of size $\lceil n/d \rceil!$. This can be compared with Construction 1 of [23] which gives a code of size $(\lceil n/d \rceil!)^{n \bmod d} (\lfloor n/d \rfloor!)^{d - (n \bmod d)}$. If we denote the rate of the code from Construction F by R , and the rate of the code from Construction 1 of [23] by R' , then

$$\frac{R}{R'} = \frac{\log_2(\lceil n/d \rceil!)}{\log_2((\lceil n/d \rceil!)^{n \bmod d} (\lfloor n/d \rfloor!)^{d - (n \bmod d)})} \geq \frac{1}{d}.$$

We now turn to provide a construction suited for $d = \Theta(n)$.

Construction G. Let $1 \leq d \leq n$ be positive integers. We recall Construction 1 from [23], of an (n, M, d) code,

$$C' = \{f \in S_n \mid f(i) \equiv i \pmod{d} \text{ for all } i\},$$

where

$$M = |C'| = (\lceil n/d \rceil!)^{n \bmod d} (\lfloor n/d \rfloor!)^{d - (n \bmod d)}.$$

Let k be the largest integer such that $k! \leq |C'|$, and let C'' be the set of all permutations over the set $\{n+1, n+2, \dots, n+k\}$. Assume

$$C' = \{f'_1, f'_2, \dots, f'_{|C'|}\} \\ C'' = \{f''_1, f''_2, \dots, f''_{k!}\}.$$

We now construct the code

$$C = \{f'_i \| f''_i \mid 1 \leq i \leq k!\},$$

where $\|$ denotes vector concatenation. □

Theorem 22. The code C from Construction G is an $[n+k, k, d]$ systematic code in the ℓ_∞ -metric.

Proof: The code is obviously of size $k!$, and by construction, the projection onto the first k coordinates gives all possible permutations exactly once. Since C' is a code with minimum distance d (see [23] for proof), the code C also has minimum distance of d in the ℓ_∞ -metric. ■

We now turn to analyze the asymptotic rate of the code from Construction G. Assume $d = \delta n$, where δ is a constant, $0 < \delta < 1$. By our choice of k , we have

$$\frac{1}{k+1} |C'| \leq |C| = |C''| = k! \leq |C'|. \quad (15)$$

Let R denote the rate of C , and R' denote the rate of C' , i.e.,

$$R = \frac{\log_2 k!}{n+k}, \\ R' = \frac{\log_2 |C'|}{n}.$$

Thus, by (15),

$$\left(1 - \frac{\log_2(k+1)}{\log_2 |C'|}\right) \frac{n}{n+k} \leq \frac{R}{R'} \leq \frac{n}{n+k}.$$

Since $1 \leq k \leq n$, while (see [23])

$$|C'| \geq 2^{(1-\delta)n}$$

we have

$$\lim_{n \rightarrow \infty} \frac{R}{R'} = \lim_{n \rightarrow \infty} \frac{n}{n+k}.$$

At this point we need to bound k , and we contend that

$$k \leq \frac{n}{\log_2 \log_2 n}.$$

Let us assume, for $k = \lceil n / \log_2 \log_2 n \rceil$, that we have

$$k! \leq |C'|.$$

We easily see that

$$|C'| \leq \left(\left\lceil \frac{n}{d} \right\rceil!\right)^d = \left(\left\lceil \frac{1}{\delta} \right\rceil!\right)^{\delta \lceil 1/\delta \rceil n} = \alpha^n$$

for some constant $\alpha > 1$.

On the other hand, we recall the well known bound

$$m! \geq \left(\frac{m}{e}\right)^m,$$

which holds for all positive integers m . Thus,

$$k! = \lceil n / \log_2 \log_2 n \rceil! \geq \left(\frac{n}{e \log_2 \log_2 n}\right)^{\frac{n}{\log_2 \log_2 n}}.$$

If indeed $k! \leq |C'|$ then necessarily

$$\left(\frac{n}{e \log_2 \log_2 n}\right)^{\frac{n}{\log_2 \log_2 n}} \leq \alpha^n,$$

and taking \log_2 of both sides gives us

$$\frac{n}{\log_2 \log_2 n} \log_2 \left(\frac{n}{e \log_2 \log_2 n}\right) \leq n \log_2 \alpha.$$

However, this last inequality certainly does not hold for large enough n . It therefore follows that indeed

$$k \leq \frac{n}{\log_2 \log_2 n}.$$

Finally,

$$\lim_{n \rightarrow \infty} \frac{R}{R'} = \lim_{n \rightarrow \infty} \frac{n}{n+k} \geq \lim_{n \rightarrow \infty} \frac{n}{n + \frac{n}{\log_2 \log_2 n}} = 1.$$

Essentially, when $d = \Theta(n)$, we constructed systematic codes with the same asymptotic rate and minimum distance as the non-systematic codes appearing in [23], which are currently the best codes known asymptotically. Furthermore, the construction we presented can work with any other non-systematic error-correcting code, provided it has an exponential size when $d = \Theta(n)$.

V. CONCLUSION

In this paper, we studied systematic error-correcting codes for rank modulation under two metrics: Kendall's τ -metric, and the ℓ_∞ -metric. In the former, we presented several constructions, explicit and algorithmic, and found the capacity of systematic codes. Efficient encoding and decoding schemes were also discussed. In the latter, two constructions were given, one of them asymptotically attaining the same rate as the best construction currently known in this metric.

Some open questions remain. In Kendall's τ -metric we still do not know, given n and d , what is the largest $[n, k, d]$ systematic code possible. We are also interested in the question of whether systematic perfect codes (or even general perfect codes) exist. In the ℓ_∞ -metric, we are still missing tight bounds, even asymptotically, on the parameters of general codes, as well as for systematic codes.

REFERENCES

- [1] A. Barg and A. Mazumdar, "Codes in permutations and error correction for rank modulation," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3158–3165, Jul. 2010.
- [2] G. W. Burr, M. J. Breitwisch, M. Franceschini, D. Garetto, K. Gopalakrishnan, B. Jackson, B. Kurdi, C. Lam, L. A. Lastras, A. Padilla, B. Rajendran, S. Raoux, and R. S. Shenoy, "Phase change memory technology," *J. of Vacuum Science and Technology*, vol. 28, no. 2, pp. 223–262, Mar. 2010.
- [3] S. Buzaglo and T. Etzion, "Perfect permutations codes with the Kendall's τ -metric," Oct. 2013. [Online]. Available: <http://arxiv.org/abs/1310.5515>
- [4] P. Cappelletti, C. Golla, P. Olivo, and E. Zanon, *Flash Memories*. Kluwer Academic Publishers, 1999.
- [5] T. Cover, "Enumerative source encoding," *IEEE Trans. Inform. Theory*, vol. 19, no. 1, pp. 73–77, Jan. 1973.
- [6] F. Farnoud, V. Skachek, and O. Milenkovic, "Error-correction in flash memories via codes in the Ulam metric," *IEEE Trans. Inform. Theory*, vol. 59, no. 5, pp. 3003–3020, May 2013.
- [7] S. W. Golomb and L. R. Welch, "Perfect codes in the Lee metric and the packing of polyominoes," *SIAM J. Appl. Math.*, vol. 18, no. 2, pp. 302–317, Jan. 1970.
- [8] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. Inform. Theory*, vol. 55, no. 6, pp. 2659–2673, Jun. 2009.
- [9] A. Jiang, M. Schwartz, and J. Bruck, "Correcting charge-constrained errors in the rank-modulation scheme," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2112–2120, May 2010.
- [10] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*. Oxford University Press, NY, 1990.
- [11] T. Kløve, "Generating functions for the number of permutations with limited displacement," *Elec. J. of Comb.*, vol. 16, pp. 1–11, 2009.
- [12] —, "Lower bounds on the size of spheres of permutations under the Chebychev distance," *Designs, Codes and Cryptography*, vol. 59, no. 1–3, pp. 183–191, 2011.
- [13] T. Kløve, T.-T. Lin, S.-C. Tsai, and W.-G. Tzeng, "Permutation arrays under the Chebyshev distance," *IEEE Trans. Inform. Theory*, vol. 56, no. 6, pp. 2611–2617, Jun. 2010.
- [14] D. E. Knuth, *The Art of Computer Programming Volume 3: Sorting and Searching*. Addison Wesley, Reading, MA, 2nd ed., 1998.
- [15] J. Leech and N. J. A. Sloane, "Sphere packings and error-correcting codes," *Can. J. Math.*, vol. XXIII, no. 4, pp. 718–745, 1971.
- [16] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North-Holland, 1978.
- [17] M. Mares and M. Straka, "Linear-time ranking of permutations," *Algorithms-ESA*, pp. 187–193, 2007.
- [18] A. Mazumdar, A. Barg, and G. Zémor, "Constructions of rank modulation codes," in *Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT2011)*, St. Petersburg, Russia, Aug. 2011, pp. 834–838.
- [19] R. M. Roth and P. H. Siegel, "Lee-metric BCH codes and their application to constrained and partial-response channels," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1083–1096, Jul. 1994.
- [20] C. D. Savage, "A survey of combinatorial Gray codes," *SIAM Rev.*, vol. 39, no. 4, pp. 605–629, Dec. 1997.
- [21] M. Schwartz, "Efficiently computing the permanent and Hafnian of some banded Toeplitz matrices," *Linear Algebra and its Applications*, vol. 430, no. 4, pp. 1364–1374, Feb. 2009.
- [22] M. Schwartz and I. Tamo, "Optimal permutation anticode with the infinity norm via permanents of $(0,1)$ -matrices," *J. Combin. Theory Ser. A*, vol. 118, pp. 1761–1774, 2011.
- [23] I. Tamo and M. Schwartz, "Correcting limited-magnitude errors in the rank-modulation scheme," *IEEE Trans. Inform. Theory*, vol. 56, no. 6, pp. 2551–2560, Jun. 2010.
- [24] —, "On the labeling problem of permutation group codes for the infinity metric," *IEEE Trans. Inform. Theory*, vol. 58, no. 10, pp. 6595–6604, Oct. 2012.